# W O R L D W I D E

# wwPDB

## PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report (i)

Aug 26, 2023 – 09:02 am BST

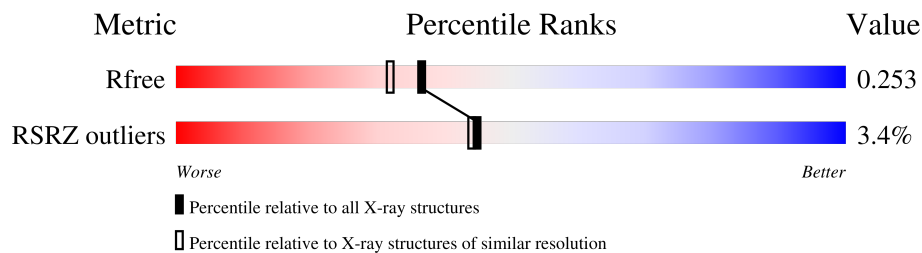| | | |
|---|---|---|
| PDB ID | : | 8BKE |
| Title | : | Crystal structure of the Klebsiella phage KP34p57 capsular depolymerase |
| Authors | : | Berisio, R.; Squeglia, F. |
| Deposited on | : | 2022-11-09 |
| Resolution | : | 2.00 Å(reported) |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

| Metric | Percentile Ranks | Value |
|--------|------------------|-------|
| Rfree | | 0.253 |
| RSRZ outliers | | 3.4% |

*Worse*                                                                *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|--------|--------------------------|----------------------------------------------------|
| $R_{free}$ | 130704 | 8085 (2.00-2.00) |
| RSRZ outliers | 127900 | 7900 (2.00-2.00) |

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

## 2   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6893 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Uncharacterized protein 57.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | AAA | 428 | Total 3237 | C 2033 | N 573 | O 620 | S 11 | 0 | 0 | 0 |
| 1 | BBB | 428 | Total 3240 | C 2034 | N 573 | O 622 | S 11 | 0 | 0 | 0 |

There are 16 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| AAA | 631 | LYS | - | expression tag | UNP D1L302 |
| AAA | 632 | GLY | - | expression tag | UNP D1L302 |
| AAA | 633 | HIS | - | expression tag | UNP D1L302 |
| AAA | 634 | HIS | - | expression tag | UNP D1L302 |
| AAA | 635 | HIS | - | expression tag | UNP D1L302 |
| AAA | 636 | HIS | - | expression tag | UNP D1L302 |
| AAA | 637 | HIS | - | expression tag | UNP D1L302 |
| AAA | 638 | HIS | - | expression tag | UNP D1L302 |
| BBB | 631 | LYS | - | expression tag | UNP D1L302 |
| BBB | 632 | GLY | - | expression tag | UNP D1L302 |
| BBB | 633 | HIS | - | expression tag | UNP D1L302 |
| BBB | 634 | HIS | - | expression tag | UNP D1L302 |
| BBB | 635 | HIS | - | expression tag | UNP D1L302 |
| BBB | 636 | HIS | - | expression tag | UNP D1L302 |
| BBB | 637 | HIS | - | expression tag | UNP D1L302 |
| BBB | 638 | HIS | - | expression tag | UNP D1L302 |

- Molecule 2 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 2 | AAA | 254 | Total 254 | O 254 | 0 | 0 |
| 2 | BBB | 162 | Total 162 | O 162 | 0 | 0 |

MolProbity failed to run properly - this section is therefore empty.

# 3 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 38.34Å    140.99Å    81.92Å<br>90.00°    102.63°    90.00° | Depositor |
| Resolution (Å) | 15.00  –  2.00<br>15.00  –  2.00 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 78.5 (15.00-2.00)<br>78.7 (15.00-2.00) | Depositor<br>EDS |
| $R_{merge}$ | 0.11 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 3.13 (at 2.00Å) | Xtriage |
| Refinement program | REFMAC 5.8.0257 | Depositor |
| R, $R_{free}$ | 0.204  ,  0.247<br>0.211  ,  0.253 | Depositor<br>DCC |
| $R_{free}$ test set | 746 reflections (1.66%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 25.4 | Xtriage |
| Anisotropy | 0.188 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.38 , 46.3 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.47$, $< L^2 > = 0.30$ | Xtriage |
| Estimated twinning fraction | 0.056 for h,-k,-h-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.94 | EDS |
| Total number of atoms | 6893 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 33.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.94% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \|L\| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

# 4 Model quality ⓘ

## 4.1 Standard geometry ⓘ

MolProbity failed to run properly - this section is therefore empty.

## 4.2 Too-close contacts ⓘ

MolProbity failed to run properly - this section is therefore empty.

## 4.3 Torsion angles ⓘ

### 4.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section is therefore empty.

### 4.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section is therefore empty.

### 4.3.3 RNA ⓘ

MolProbity failed to run properly - this section is therefore empty.

## 4.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 4.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 4.7 Other polymers ⓘ

There are no such residues in this entry.

## 4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 5 Fit of model and data ⓘ

## 5.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, $95^{th}$ percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|-----|-------|----------|--------|---------|----|----|-------------|-------|
| 1 | AAA | 428/638 (67%) | -0.14 | 10 (2%) | 60 | 59 | 18, 28, 42, 71 | 6 (1%) |
| 1 | BBB | 428/638 (67%) | 0.15 | 19 (4%) | 34 | 33 | 21, 35, 53, 77 | 2 (0%) |
| All | All | 856/1276 (67%) | 0.00 | 29 (3%) | 45 | 44 | 18, 31, 50, 77 | 8 (0%) |

The worst 5 of 29 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | BBB | 27 | VAL | 5.8 |
| 1 | BBB | 151 | ASP | 5.6 |
| 1 | AAA | 454 | LEU | 5.2 |
| 1 | BBB | 28 | PHE | 4.3 |
| 1 | BBB | 286 | GLN | 3.6 |

## 5.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.4 Ligands ⓘ

There are no ligands in this entry.

## 5.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK